## 0.1 Multilayer Feedforward Networks are Universal Approximators

In 1988 Maxwell Stinchcombe and Halber White show that Multilayer Feedforward Networks with one hidden layer are Universal Approximators.

The proof is based on that arbitrary squashing function are capable of approximating any Borel measurable function from one finite dimensional space to another to any desired degree of accuracy, provided sufficiently hidden units are available. In this sense Feedforward networks are a class of Universal approximators. (REWRITE THREE LAST SENTENCES)

Firstly we are going to start with some basic definitions.

**Definition 0.1.1.** Affine functions

Affine functions are defined as all the function of the form $A(x) = w \cdot x + b$ from $\mathbb{R}^d$ to $\mathbb{R}$ where $d \in \mathbb{N} - \{0\}$, $w$ and $x$ are vector in $\mathbb{R}^d$, $\cdot$ denotes the usual dot product of vector, and $b \in \mathbb{R}$ is a scalar.